音声ポータル実証実験の最終報告

2004/05/24

VoiceXML部会
アプリケーション検討WG
渡辺 聡

音声ポータル

- 音声ポータルとは
  - 例）「インターネット上の様々な情報を、電話で取り出すサービスのこと」
  - アプリケーション検討WGでは．．．
    - 「音声インタフェースを使った情報配信型のサービス」と広く定義
- サービス状況
  - 2000年頃から、米でサービス開始
  - 2001年から国内でもサービス開始

XML Consortium

VoiceXMLの利用イメージ



電話網

インターネット

VXML
実行環境
音声合成
音声認識

VoiceXML

VoiceXML
CGI
DB

音声ブラウザ

コンテンツサーバ
(WWWサーバ)

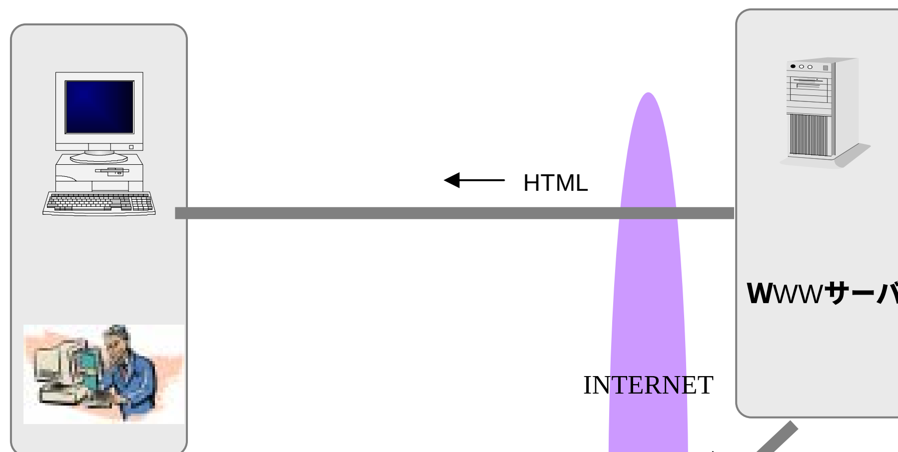
実験の概要



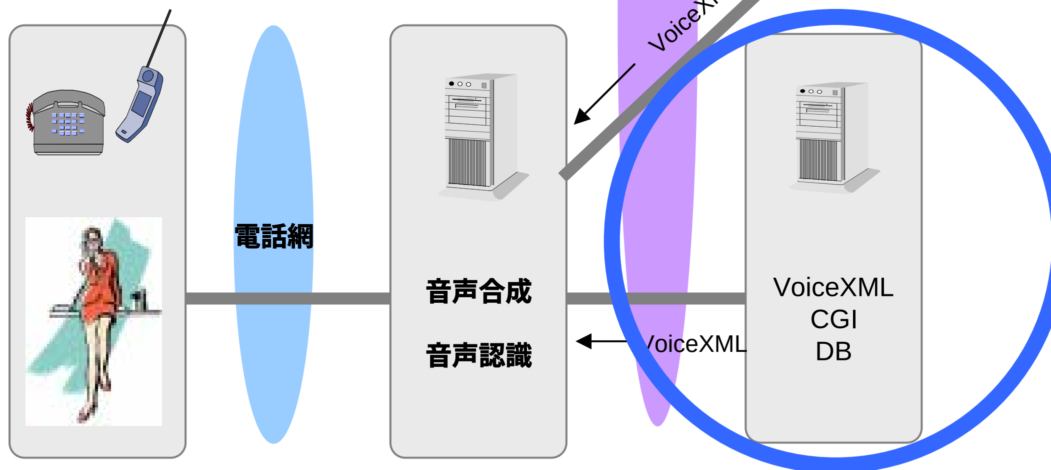
- [実験の背景]
  - VoiceXMLを用いた音声による情報配信サービス(音声ポータル)の普及促進。
- [実験の目的]
  - ユーザインタフェース（UI）知見の実践的検証
    - 音声ポータル簡易評価方法
    - 音声UIの典型的な問題
  - システム構築プロセスを共有
  - 実際に動いているVoiceXMLシステムを、広く紹介・意見交換
- [活動内容]
  - 電話を使った資料検索請求サービス等の実験アプリケーションを構築
  - 参加企業が3チームに分かれて、共同開発
    - 実験企画チーム ・ コンテンツ作成チーム ・技術検討チーム
  - 活動を3つのターンで分けて、それぞれでユーザテスト等により評価・改善
- [参加企業]
  - MITシステム研究所、沖電気工業、しーこむ、都築ソフトウェア、トッパン・フォームズ、日本テレコム、日立製作所、富士通、富士通研究所 (50音順)

XML Consortium

システム構成



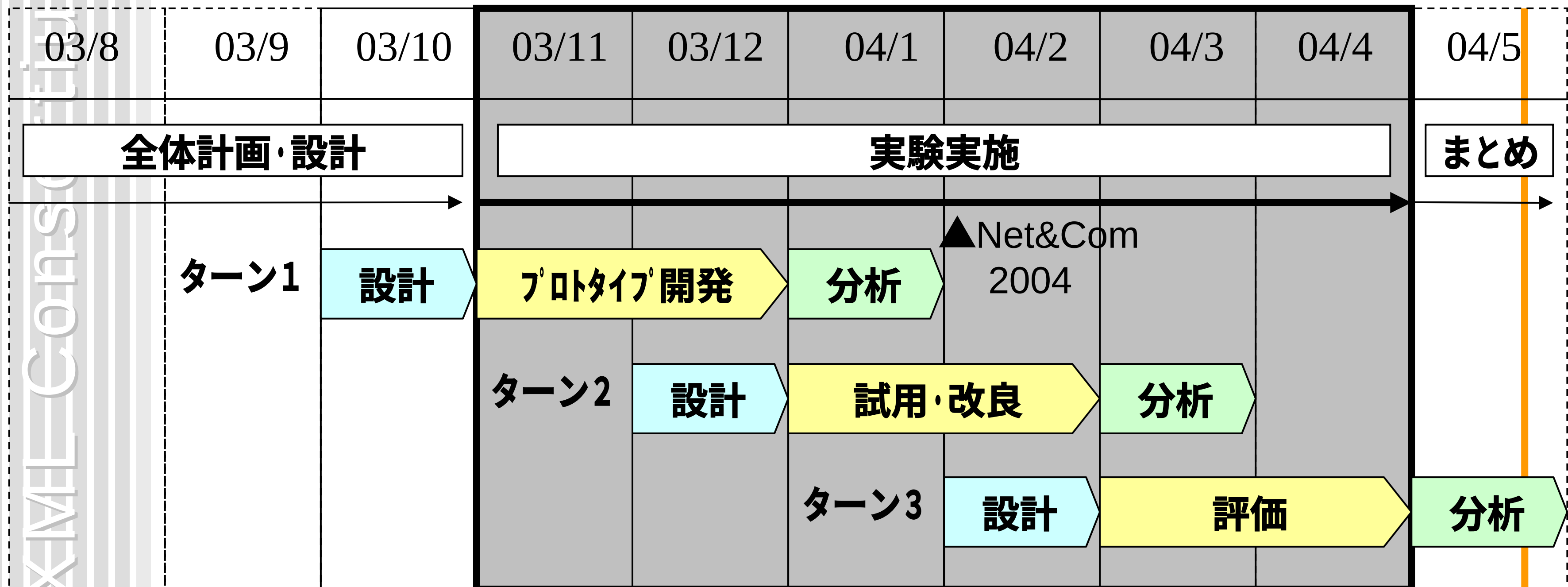
通常のWebアクセス
HTML
INTERNET
WWWサーバ



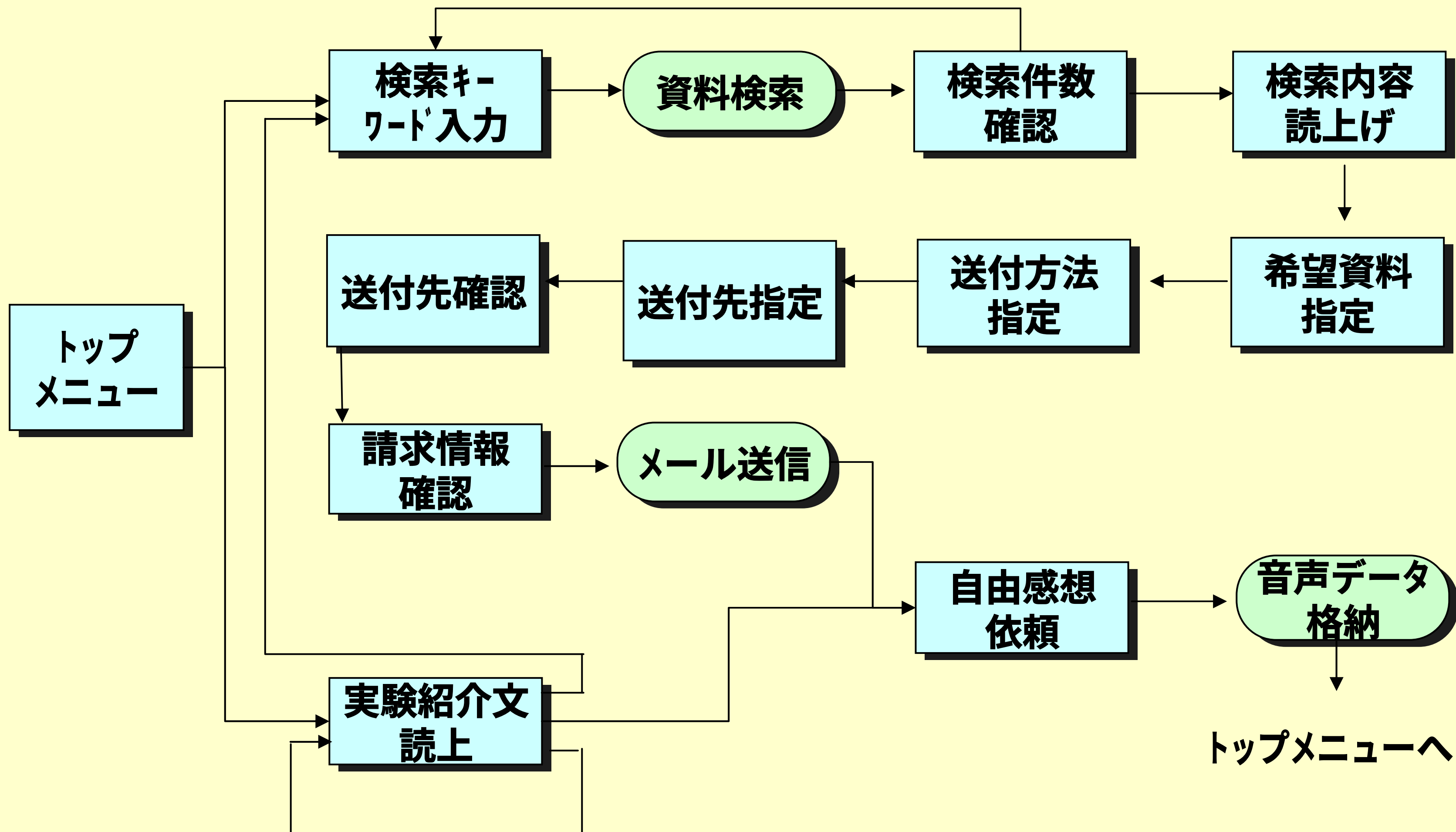
音声を使った
Webアクセス
電話網
VoiceXML
音声合成
音声認識
VoiceXML
CGI
DB
音声プラットフォーム
ボイスブラウザ
コンテンツサーバ
WWWサーバ

スケジュール

- 期間
  - 03/11~04/04(6ヶ月間)。
  - 3つのターンに区切って実施

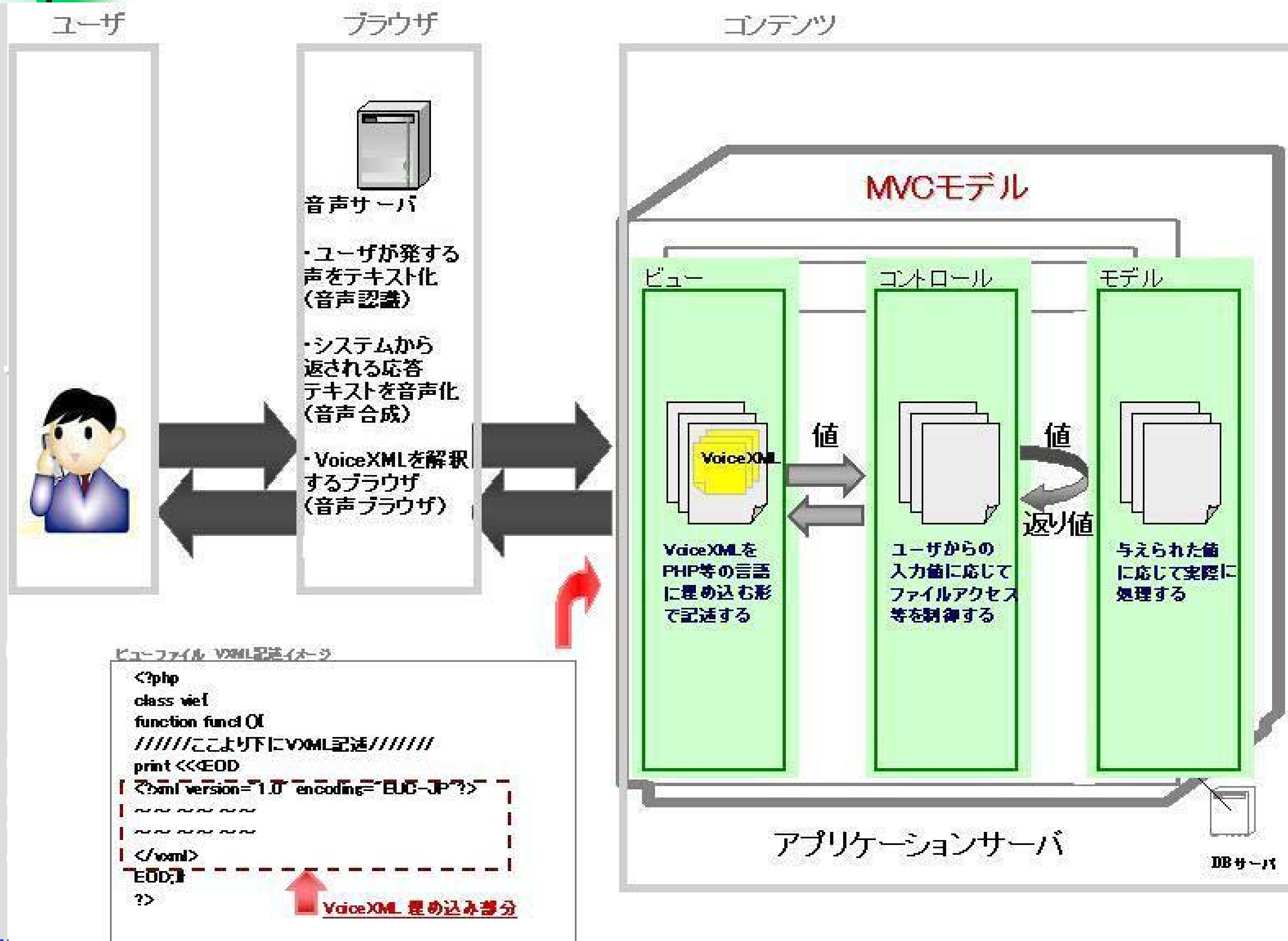


コールフロー



検索ｷｰ ﾜｰﾄﾞ入力
資料検索
検索件数 確認
検索内容 読上げ
トップ メニュー
送付先確認
送付先指定
送付方法 指定
希望資料 指定
請求情報 確認
メール送信
自由感想 依頼
音声データ 格納
実験紹介文 読上
トップメニューへ

VoiceXML部の実装方法





データ構成

AUTHOR
- ◎AUTHOR_ID
- AUTHOR
- AUTHOR_SEI_PRON
- AUTHOR_MEI_PRON
- COMPANY
- COMPANY_PRON

DOCUMENT
- ◎DOC_ID
- TITLE
- TITLE_PRON
- SHORT_TITLE
- SHORT_TITLE_PRON
- DATE
- AUTHOR_ID
- AUTHOR_ID2
- AUTHOR_ID3
- AUTHOR_ID4
- ○PATH

DOC_KEYWORD_MAP
- ◎DOC_ID
- ◎WORD_ID

KEYWORD
- ◎PRON_ID
- WORD_ID
- WORD
- ○WORD_PRON
- ATTRIBUTE

USERS
- ◎USER_ID
- NAME
- NAME_SEI_PRON
- NAME_MEI_PRON
- COMPANY
- COMPANY_PRON
- ZIPCODE
- ADDRESS
- ADDRESS_PRON
- TELNO
- EMAIL
- EMAIL_PRON
- EMAIL_MOBILE
- EMAIL_MOBILE_PRON

※図中下線はﾃｰﾌﾞﾙ名を表し、◎、○及び●はそれぞれ主ｷｰ、一意ｷｰ及び外部ｷｰを表す

プロンプトの設計



Barge in

着信

こちらはVoice XML部会の音声実験ポータルです。
サービスを受ける前に、実験の位置付け、目的などをお聞きになるには、目的、とおっしゃってください。お聞きにならない場合は、つぎへ、とおっしゃってください。

情報提供

・・・・実験紹介は以上です。では、資料請求サービスをお受けになるかたは、「つぎへ」とおっしゃってください。また、終了されるかたは「おわり」とおっしゃって、電話をお切り下さい。

「おわり」 → 切断

「次へ」 → (資料請求へ)

資料請求

こちらは、資料請求サービスです。Voice XML部会の技術資料、議事録を検索、そして必要な資料はお送りいたします。はじめに、検索したい資料のキーワードを3個までおっしゃってさい。

「キーワード3個まで」

キーワード確認

キーワード検索

検索結果0件

キーワードに該当する資料がありませんでした。検索をやり直す場合は、「検索はじめ」、ひとつ戻る場合は「ひとつもどる」とおっしゃってください。

「終了」 → 切断or TOP

「検索始め」 → (資料請求へ)

検索結果1件

検索結果は 1件です。

→ (検索結果読上へ)

検索結果複数

検索結果は ○△ 件です。
結果を確認しますか？
確認したい場合は、「リスト」
結果をさらに絞り込みたい場合は、「絞込む」
前の結果に戻りたい場合は、「ひとつ戻る」
とおっしゃってください。
では、どうぞ。

「リスト」 → 検索結果読上げ

「ひとつ戻る 絞込む」 → (資料請求へ)

検索結果読上げ

あなたの検索した資料は・・・
こちらで資料を確定する場合は、「リスト」と、
この中からご希望の資料をお選びになる場合は、希望する資料の番号をおっしゃってください。
あるいは、ひとつ前の検索に戻る場合は、「ひとつ戻る」
さらに絞込み検索を続ける場合は、「絞り込む」
とおっしゃってください。では、どうぞ。

「ひとつ戻る 絞込む」 → (資料請求へ)

「資料番号 (3個まで)」「リスト」 → (送付資料決定へ)

プロトタイプの問題点



問題	対処

■**検索部**

- キーワードをうまく言えない　⇒ガイダンスの表現を統一
  - 音声認識要求時/失敗時
- 1ターンをシンプルに　⇒検索タスクに主要経路

■**ユーザ確定部**

- 多様な入力を受け付けてほしい　⇒認証失敗の無限ループ回避

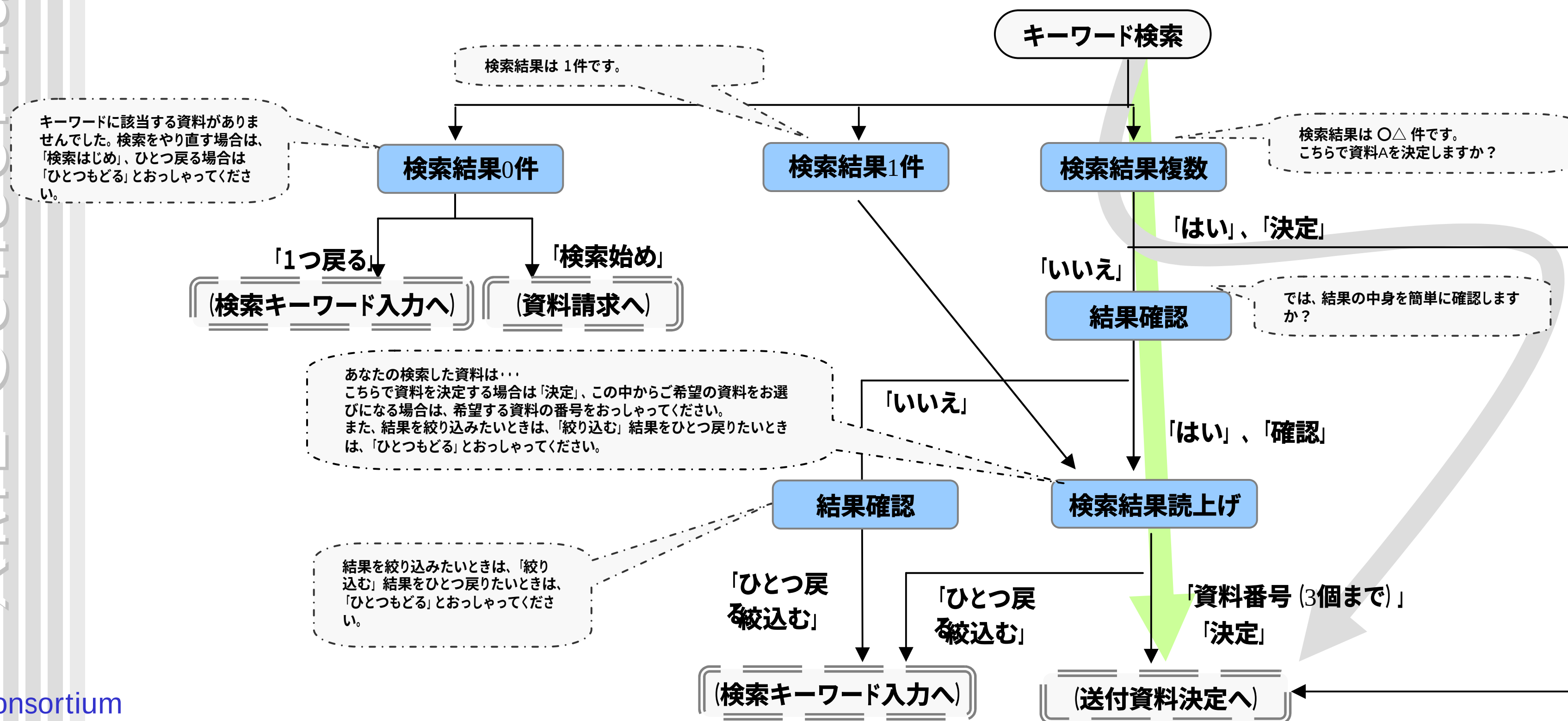
■**メッセージ確認部**

- とまどう。メッセージを上手く残せない　⇒ユーザになじみの深い留守番電話の録音を模倣

対処例：「検索タスクに主要路」

仮説　：　ユーザは一回の検索でＯＫとなることが多い
⇒　　　ユーザに極力負担のかからない経路を誘導するフロー設計

※一番シンプルな経路にユーザを導く（キーワード入力→（結果確認）→資料決定）
※極力、特別な言葉を使わせない　（「はい」「いいえ」「決定」など）
※ユーザに一度に多くの選択肢を与えない



対処例：メッセージ録音の改善



- 問題：メッセージをとっさにまとめられない。
  - 話す内容をある程度限定
    - 「感想を」⇒「使いづらい点を」
  - メッセージ録音する旨を、事前に通知(未実装)
- 問題：録音処理がうまくいかない
  - 日常使い慣れた手続きを模倣　(留守番電話)
    - 始まりきっかけは、ビープ音
    - 録音終了きっかけはプッシュフォン

XML Consortium

典型的問題(概要)

XML Consortium

■プロンプトに関するもの
・注意して聞いておかないといけない所をわかるようにしてほしい
・プロンプトとして選択肢を示す場合は、簡潔に示してほしい
・プロンプト(列挙/非列挙) に対する発声揺らぎは、吸収して認識してほしい。

■フィードバックに関するもの
・システム状態・ユーザ可能動作 (発話内容/タイミング) を明確にしてほしい
・音声認識がうまく働かないときには、状況に加えて、「どうすべきか」を教えてほしい
・重要な処理については確認フェーズがほしい
・フィードバックに効果音を導入してほしい

■基本操作に関するもの
・多くのコンテンツで共通な操作 (基本操作) は、なるべく広い範囲で言葉を統一してほしい
・基本操作は、ボタンで代用できるようにしてほしい
・上記操作 (正式名称や代用ボタン) を視覚的に確認できる手段がほしい。

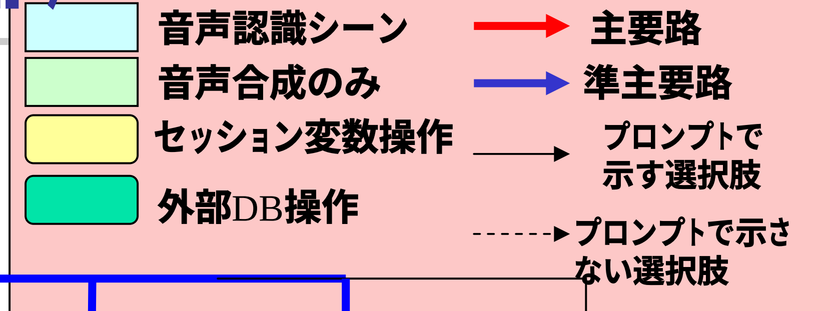
■その他(未分類)
・自分のおかれている状況 (位置) をわかるようにしてほしい。
・考え中にプロンプトを出さないでほしい。

出展後の改良

- ユーザの感想
  - 細かい問題はあるが、概ね目的は達成できそうなｼｽﾃﾑである。
- 改良点
  - 主要路の変更　：　検索結果による条件分岐
    - 6件以上：ｱｸｼｮﾝ選択、1-5件：自動ﾀｲﾄﾙ読上げ、0件ｱﾝﾄﾞｩｳ
  - キーワードの認識精度低下に対してはユーザの負荷を出来るだけかけない形で認識結果を確認/確定させる処理を追加
  - 読み上げる情報量を減らす
    - (旧)「～タイトル～作成者～」　⇒　(新)「～タイトル～」
  - 音量、テンポの修正
    - prosタグの修正により、音量→UP、テンポ→ゆっくり
  - 音声認識タイミング時の効果音の利用
  - ユーザの意見を反映したプロンプト検討

XML Consortium

コールフロー (検索部)



着信
「資料請求」
キーワード入力
「やり直し」
「W3C」
入力確認「W3Cで検索します」
資料検索
0件
検索対象アンドゥ
検索対象保持
検索対象初期化
2~5件
6件以上
1件
「一つ戻る」
「絞込み」
「検索はじめ」
結果 (件数) 確認 (内容を読上げますか？)
アクション確認 (結果を絞込みたい時は、…)
「いいえ」
「はい」
Short_title読上げ
「いいえ」
「いいえ」
資料決定 (こちらで決定しますか)
「はい」
「決定」、番号による指定

評価方法

1. 利用シナリオの想定

どんなユーザが、どんな目的で
ＶＰを利用するかを想定する

2. 試用タスクの決定

この調査では、ＶＰに電話をして、
具体的に何をしてもらうかを決定する

3. 試用者の確保

上記1.、2. にふさわしい試用者を確保する。

4. 試用（タスクの実行）

ＶＰを、タスクに沿って試用してもらう。
操作中、使い勝手についての感想を
口走ってもらい、観察者がメモをとる。

5. 問題点の整理（ｲﾝﾀﾋﾞｭｰ）

観察者は、幾つかの視点から、
「使い勝手」についてインタビューし、
問題点を整理する。

タスク

XML Consortium

- タスク1
  - 「W3C」に関する資料を取り寄せてください（取り寄せた資料）
- タスク2
  - 「VoiceXML」の「仕様」に関する資料を取り寄せてください
- タスク3
  - 下記（リスト)の目次を持つ文書(1ファイル）を、適切な検索キーワードを用いて探し出してください
    - 組み込み系音声システムとは？
    - 音声技術の現状の限界および今後の課題
    - 組込み機器のマルチメディア対応音声システムについて
    - 組み込み機器へのVoiceXMLの適用について
    - 組み込み用音声認識/合成エンジンの紹介
    - 組み込み系音声システムの製品紹介
    - 組み込み系音声システムの現状と今後の展開について

評価項目

XML Consortium

- 本評価で利用した10項目 (参考：ニールセンの10項目)
  - ①各ガイダンスの中で、不要と感じた情報、不自然と感じた情報はありましたか？
  - ②意味のよくわからない言葉や表現がありましたか？
  - ③対話を進めていく中で、以前に聞いた情報を覚えていなかったためにうまく話せないところがありましたか？
  - ④同じ言葉で違う事を意味しているように聞こえたりなどで、ガイダンスを聞いていて混乱する事はありましたか？
  - ⑤システムの反応がない/システムの作業状況が分からなくて、何をすればよいか戸惑ったりする事はありましたか？
  - ⑥「エラー」or「意図しない終わり方」をしましたか？あれば、箇所を教えてください。それは何で起きたと考えられますか？
  - ⑦意図しない場所に移動してしまい、「脱出」をこころみたシーンがありましたか？その時、非常口はうまく見つかりましたか？
  - ⑧はじめて使った際に、何をすればよいか分からなくなったりなど、ガイダンスで説明不足だと感じる事がありましたか？
  - ⑨システムが要求する各コマンドの意味はすぐ分かりましたか？
  - ⑩提供されたマニュアル（取扱説明書）は明確で要点を得たモノでしたか？

評価結果

XML Consortium

項目	初心者	熟練者
1．不要･不自然な情報を入れない	3.0	4.3
2．意味不明の表現を使わない	2.7	4.7
3．記憶の負荷を減らす	3.7	3.3
4．混乱するガイダンスを入れない	3.7	4.3
5．システム状態を通知する	3.0	3.3
6．意図に反するエラーや遷移防止	2.3	4.0
7．脱出時の非常口明確化	2.7	4.0
8．十分な説明	4.7	4.0
9．コマンド語の自明性	3.7	4.3
10．マニュアルの判り易さ	3.0	4.3
平均	3.3	4.1

5＝全く問題ない、
4＝問題はあるが許容範囲、
3＝問題はあり，
目的達成可能だが，
修正が好ましい、
2＝問題あり、
修正しないと、
目的達成不能、
1＝修正以前の問題である。

※目的達成可能 ：
本システムでは「消耗の資料を手に入れられる」こと

本システムは「目的達成可能」なシステムである。
ただし、特に初心者に対しては、いくつかの改良が好ましい。

改善が好ましいポイント

- 初心者に対して
  - 誤認識等による、意図せぬ遷移等を極力起こさせない
    - 音声検出タイミングの最適化
    - 連続音声認識のショートカット化
    - キーワード数の拡充
  - 音声合成音がききにくい
    - 話速の再度みなおし(一貫性の視点から)
    - 録音音声の利用
- 熟練者に対して
  - 記憶負荷の低減(検索キーワード,検索結果)
    - 検索状態情報の問合せ手続きの定式化
    - バスケット機能の付加

XML Consortium

蓄積ノウハウの検証

- 典型的問題点
  - 概ねユーザコメントとを網羅していた。：妥当性確認
  - 特に、ユーザコメントを咀嚼する際、有効性を発揮
  - ただし、一般的問題点であるため、アプリケーション依存で対処方法が異なる。利用方法に注意が必要。
- 簡易評価方法
  - システムのUI的弱点を抽出できた。：有効性確認
  - 経験のない評価者でも実施でき、リーズナブル（時間面、コスト面）な方法であることを確認。
  - 詳細、特に各項目の具体化については今後の課題。
  - またタスク決定に関するのガイドラインは別途必要。

XML Consortium

実験まとめ

- WG活動で蓄積してきたUI知見をベースに、「資料請求」システムを構築した。
- 同様に蓄積してきた評価方法について、システムのUI評価を行った。
- その結果、構築したシステムは「目的達成可能」であることがわかった。
- また、評価によって、初心者・熟練者に対してさらに改良すべき箇所を抽出できた。
- 「典型的問題点」については妥当性を、「簡易評価方法」については有効性を、それぞれ確認できた。

Net&Com2004　来訪者像



0 5 10 15 20 25 30

製造業
商社/卸/小売業
サービス業
金融/証券/保険
建設/不動産
学校/教育機関
運輸
病院/医療機関
印刷/出版/放送/広告
農林/水産/鉱業
官公庁/公共機関・団体
学生
その他
システムインテグレータ
コンサルタント
VAR/ディーラー/リセラ
通信サービスプロバイダ/電気通信事業者
その他のインターネットビジネス
システムハウス/ソフトハウス
コンピュータメーカー/周辺機器メーカー
ネットワーク・通信関連機器メーカー

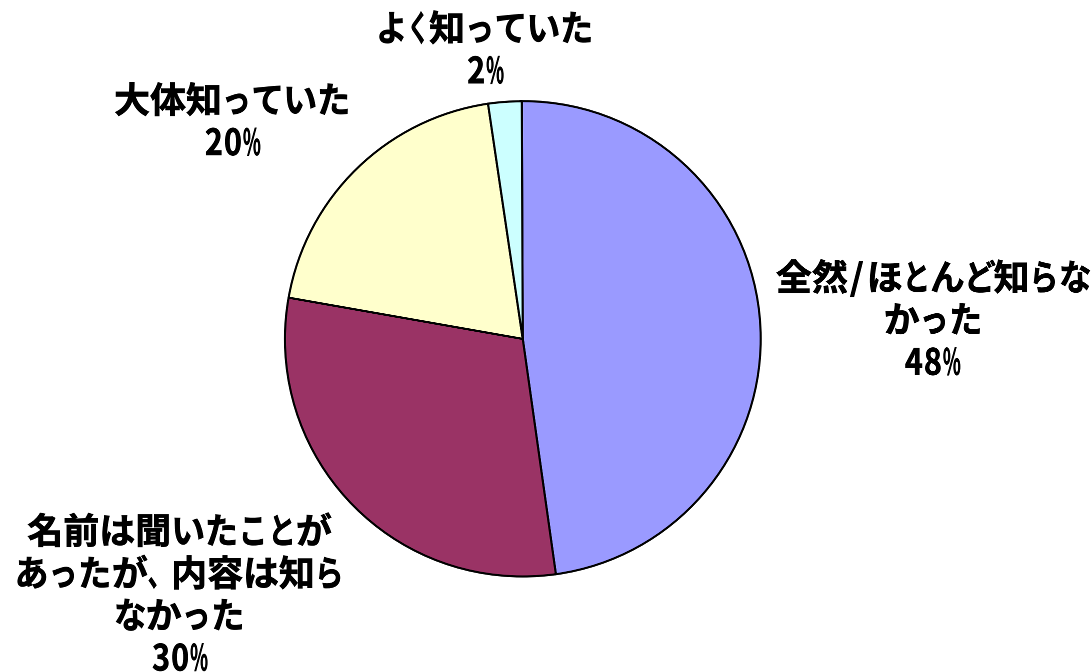
52

専門セミナー事前登録者より。ユーザ系43%、**ベンダ系**57%

来訪者のプロフィール（1）



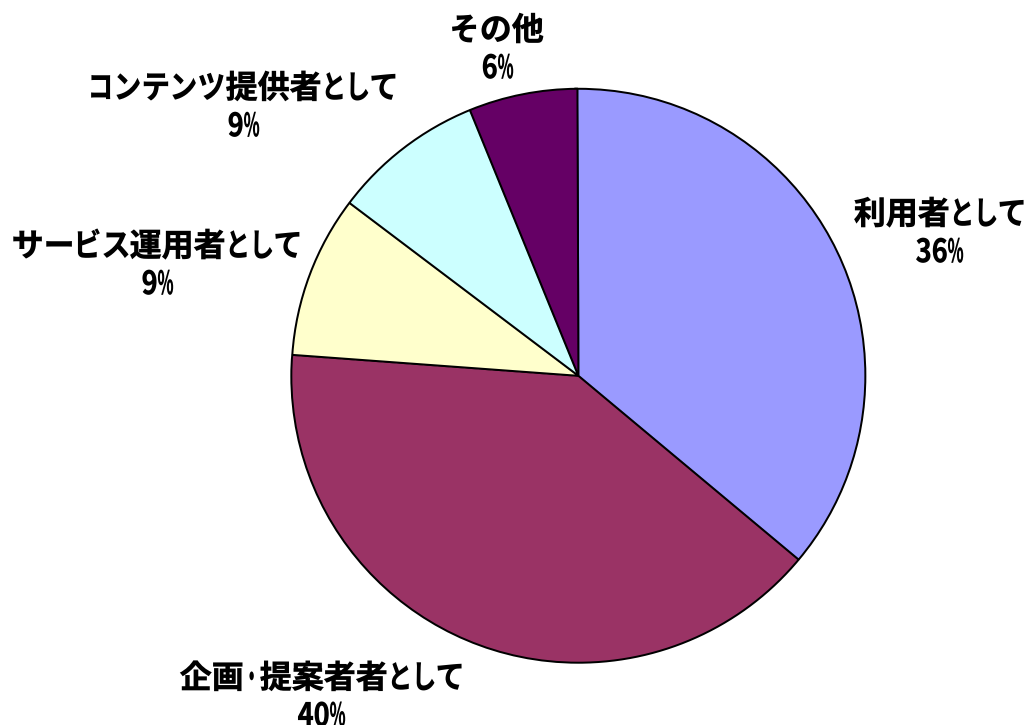
Q.「音声ポータル」/VoiceXMLをご存知でしたか？



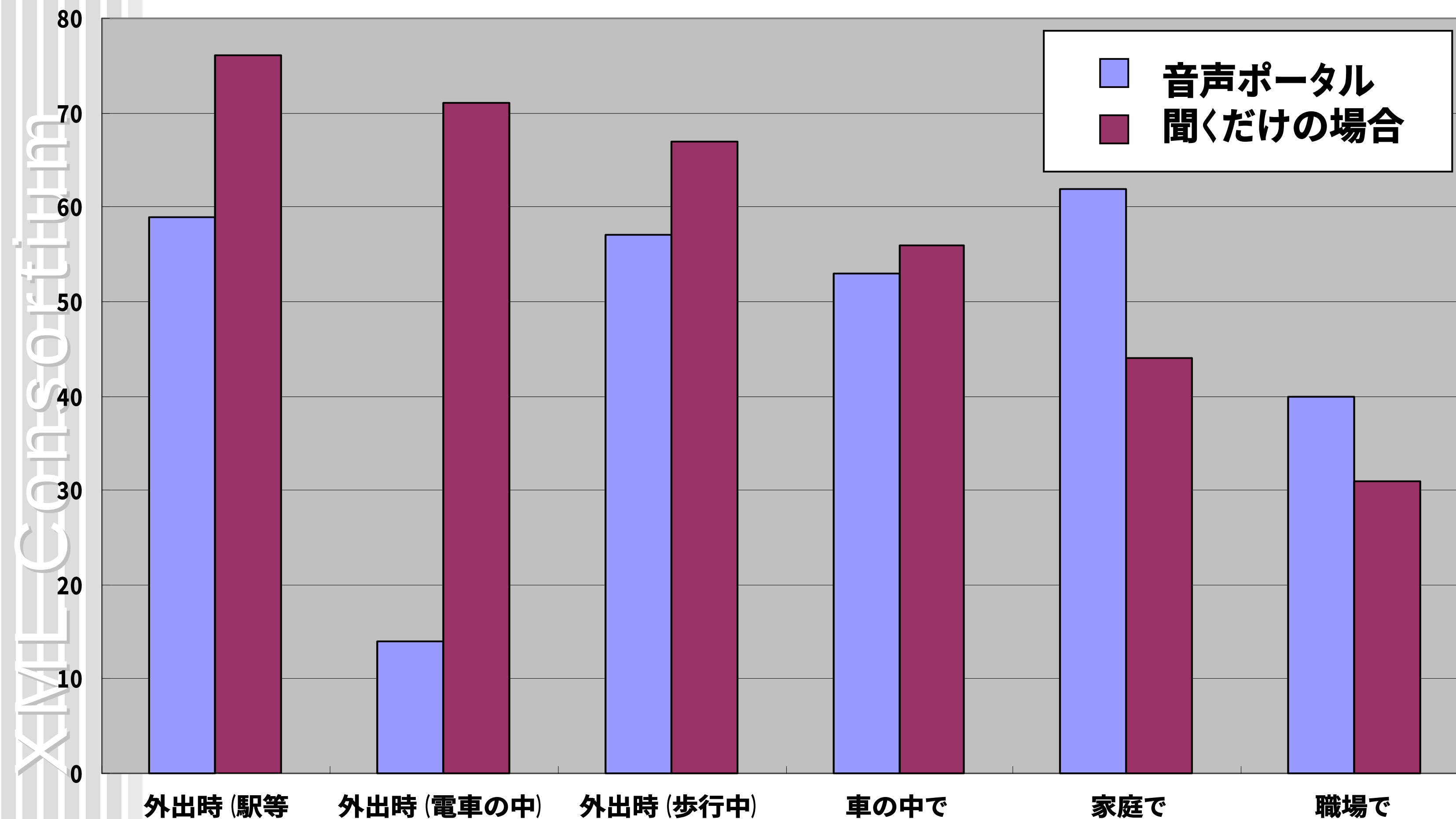
XML Consortium

来訪者のプロフィール (2)

Q: あなたが「音声ポータル」に関わる場合、どのような立場になりますか？



利用したいシーン



Q. 利用したいシーンは？
80
70
60
50
40
30
20
10
0
音声ポータル
聞くだけの場合
外出時 (駅等
外出時 (電車の中)
外出時 (歩行中)
車の中で
家庭で
職場で

音声に向く/向かないｻｰﾋﾞｽ



Q.　音声ポータルに向くと思うもの、向かないと思うものを教えてください

向くと思うもの
向かないと思うもの

70
60
50
40
30
20
10
0

見積もりｼｽﾃﾑ
社内DBｱｸｾｽ
電子ﾒｰﾙ読上げ
在庫管理
株価・残高
ニュース速報
乗り換え案内
道案内
チケット予約
ﾚｽﾄﾗﾝ案内
資料・ｶﾀﾛｸﾞ請求
ｻﾎﾟｰﾄｾﾝﾀ
占い・診断
クイズ・雑学
ゲーム
英語学習

音声ポータルのメリット



Q.「音声ポータル」のよいところはどこだと思いますか？

対話という自然な手段でｱｸｾｽできること	58
目が疲れないこと	21
「ながら利用」できること	47
ある意味面白いこと･楽しいこと	11
気楽に相槌をうっていれば情報が得れること	7
文字を打ち込まなくてもよいこと	67
電話という身近な装置でアクセスできること	62
その他	6

0 10 20 30 40 50 60 70 80

活動成果

- VoiceXMLを使った目的達成可能な「資料請求」システムを実現
- UI知見の検証
  - 「典型的問題点」の妥当性確認
  - 「評価方法」の有効性確認
- システム構築プロセスの共有
  - メンバ間で役割を分担し実行
  - 特に、改良・評価の手順を具体化
- 音声ポータル/VoiceXMLの普及促進
  - 各種セミナ/展示会等での直接対話による紹介
  - メールマガジン等で呼びかけ、オンラインでの紹介

今後に向けて

- 得られたノウハウを開発者･企画者向け書籍としてまとめ、VoiceXML/音声ポータルの、さらなる普及を図る。
  - UI知見、開発プロセス等
- 得られた知見に基づき、標準化(W3C等)への要求を検討する。
- 他のXMLシステム･ユビキタスシステムとの連携の中で、新たな音声応用システムを形を実現する。

XML Consortium

活動を振り返って

- 専門分野や立場の異なるメンバで、一つの実験を遂行できたのは有意義であった。
- 多くのユーザの方と、システム試用を通してディスカッションをできたのは貴重であった。
  - 音声アプリケーションのニーズを改めて確認できた。
- 今回の成果を更に展開し、音声応用システムの市場を広げていく。